Cognitive and Noncognitive Improvements Among Cadets at the Washington Youth Academy (WYA)

Warren T. Sutton, Lauren D. Malone, and Jennifer R. Atkin

December 2014

CNA
ANALYSIS & SOLUTIONS

This document contains the best opinion of CNA at the time of issue.
It does not necessarily represent the opinion of the sponsor.

Distribution

Distribution unlimited. Specific authority: N00014-11-D-0323.

Copies of this document can be obtained through the Defense Technical Information Center at www.dtic.mil or contact CNA Document Control and Distribution Section at 703-824-2123.

Photography Credit: Photo taken from the Washington Youth Academy website.

Approved by: **December 2014**

David M. Rodney, Director
Fleet and Operational Manpower Team
Resource Analysis Division

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	**2. REPORT TYPE**	**3. DATES COVERED** *(From - To)*
12-2014	Final	

4. TITLE AND SUBTITLE
Cognitive and Noncognitive Improvements Among Cadets at the Washington Youth Academy (WYA)

5a. CONTRACT NUMBER
N00014-11-D-0323

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
N/A

6. AUTHOR(S)
Warren T. Sutton, Lauren D. Malone, Jennifer R. Atkin

5d. PROJECT NUMBER
N/A

5e. TASK NUMBER
A30530

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Center for Naval Analyses
3003 Washington Blvd
Arlington, VA 22201

8. PERFORMING ORGANIZATION REPORT NUMBER
DRM-2014-U-008846-Final

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DASD for Reserve Affairs (Resources)
1500 Defense Pentagon
Washington, DC 20301-1500

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This study follows CNA's 2013 analysis of changes in a class of cadets at the Washington Youth Academy (WYA) National Guard Youth ChalleNGe Program (ChalleNGe). It analyzes data from a second class of cadets and draws conclusions regarding how participation in ChalleNGe affects youths' cognitive and noncognitive growth. It also looks at the relationship between cognitive and noncognitive measures and the predictive power of noncognitive skills. Our findings suggest that the WYA ChalleNGe program has a substantial impact on cadets' noncognitive skills; however, we found no noncognitive measures that strongly predict program completion. We found statistically significant improvements in four cognitive measures. Regarding the relationship between noncognitive and cognitive growth, we found that initial math efficacy is much more important in predicting final math scores for those with low scores at the end of ChalleNGe than for those with higher scores. We also found several gender differences.

15. SUBJECT TERMS
ChalleNGe, Cadets, Cognitive, Noncognitive, Test Scores, Washington Youth Academy, Test of Adult Basic Education, High School Dropouts, Grit, Locus-of-Control

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
					Knowledge Center/Robert Richards
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			**19b. TELEPHONE NUMBER** *(include area code)*
U	U	U	SAR	44	703-824-2104

Abstract

This study follows CNA's 2013 analysis of changes in a class of cadets at the Washington Youth Academy (WYA) National Guard Youth ChalleNGe Program (ChalleNGe). It analyzes data from a second class of cadets and draws conclusions regarding how participation in ChalleNGe affects youths' cognitive and noncognitive growth. It also looks at the relationship between cognitive and noncognitive measures and the predictive power of noncognitive skills. Our findings suggest that the WYA ChalleNGe program has a substantial impact on cadets' noncognitive skills; however, we found no noncognitive measures that strongly predict program completion. We found statistically significant improvements in four cognitive measures. Regarding the relationship between noncognitive and cognitive growth, we found that initial math efficacy is much more important in predicting final math scores for those with low scores at the end of ChalleNGe than for those with higher scores. We also found several gender differences.

This page intentionally left blank.

Executive Summary

This study presents follow-on work involving the analysis of data on both cognitive (e.g., math and language arts) and noncognitive (e.g., ability to follow directions and determination/grit) changes in youth resulting from their participation in the National Guard Youth ChalleNGe Program (ChalleNGe). CNA's first analysis [1] looked at changes in one cadet class at the ChalleNGe program in Washington State, the Washington Youth Academy (WYA). In this study, we analyze data from a second class of WYA cadets and draw conclusions about how participation in ChalleNGe affects youths' cognitive and noncognitive growth. It analyzes the relationship between cognitive and noncognitive measures and the predictive power of noncognitive skills.

In this analysis, we use several sources of WYA-provided data. First, the program collected cadets' scores on the Test of Adult Basic Education (TABE) at the beginning and end of the program. Our analysis relies on the four TABE subsections, or subtests (Math Computation, Applied Math, Reading, and Language), as well as on the total score (formed by averaging subtest scores). We also use data from a survey that was designed to measure noncognitive skills. It gathered data on five measures: grit, locus-of-control, math and science efficacy, time preference, and following directions. Cadets completed the survey twice—once at the conclusion of the initial two weeks (known as pre-ChalleNGe) and again (for those that completed the program) during the last week of classes.

Our analysis first focuses on survey results regarding cadets' noncognitive skills. Specifically, we provide descriptive statistics and explain how they change over the course of the program. We then examine the progress they made in cognitive skills and analyze the relationship between noncognitive skills and program outcomes. Finally, we use the TABE data to determine if initial TABE scores, in addition to initial noncognitive scores, can be used to predict ChalleNGe program completion and how noncognitive skills influence TABE scores.

Our findings suggest that the WYA ChalleNGe program continues to have a substantial impact on cadets' noncognitive skills. The only noncognitive measure that does not improve significantly is science efficacy. The reason, however, appears to be that over 50 percent of cadets start with a science efficacy equal to 3 or above on a fixed scale of 1 to 5.

We also find that statistically significant improvements were made in all four TABE subtests that we analyzed. Cadets improved, on average, by at least 2 grade levels in every area. For the Reading subtest, this particular cadet class experienced average improvements of 3.5 grade levels, which is twice the average improvement from 2009 to 2013. In addition, we also find that there are no significant differences between the initial TABE scores of all cadets and the scores of those cadets who eventually graduate, indicating that TABE scores are not a particularly good predictor of program completion.

In fact, we are unable to find any cognitive or noncognitive measures that strongly predict program completion. We were, however, able to establish a relationship between math efficacy, initial math score, and the final math score. We determined that both the initial math score and initial math efficacy are strong predictors of the final math score, for both male and female cadets, for those with initial math scores below 6.0. Yet, for those cadets with initial math scores greater than or equal to 6.0, only the initial math score is a significant predictor of the final math score, and this finding holds for male cadets only. Thus, there will be some subpopulations for which the gains from enhanced noncognitive skills are greater than they are for others.

We also found other notable gender differences. First, male cadets have higher initial math and science efficacy, greater initial locus-of-control, and higher initial TABE math scores (in both math computation and applied math) than female cadets. Second, their scores also are statistically significantly different at the end of ChalleNGe. In terms of noncognitive skills, male graduates have higher math efficacy, science efficacy, and grit than their female counterparts. In terms of final cognitive skills, male students have statistically higher final TABE scores in the Math Computation, Applied Math, and Language subtests.

As indicated, our results on the noncognitive measures mirror several findings from the first CNA study. First, among the cadets who ultimately graduated from ChalleNGe, noncognitive skills improved on average in both studies. Cadets who finished the program had statistically significantly higher scores in all noncognitive measures, except science efficacy, where no statistically significant progress was found in our most recent work. Second, both efforts found that initial measures of noncognitive skills are not particularly good predictors of which cadets will complete the ChalleNGe program. Both studies also found gender differences in the noncognitive measures. Specifically, female cadets began the program with lower measures of efficacy (in science and math) and were less internal than male cadets. Finally, with respect to cognitive changes, both studies concluded that the TABE reading score had explanatory power over program completion. Cadets with lower reading scores were less likely to complete the program.

We suspect that the positive changes that occur in cadets' cognitive and noncognitive skills will be long lasting, but we have no quantitative metrics at this time to

determine if this is true. We therefore suggest that WYA, and other ChalleNGe academies, consider a longitudinal study of cadet performance that would allow cadets to be tracked as they reintegrate into their home high schools (where applicable) and eventually into their postsecondary schooling and career fields. As a result of the gender differences we found, we also suggest that ChalleNGe programs consider gender-tailored approaches to their curricula.

This page intentionally left blank.

Contents

This page intentionally left blank.

List of Figures

This page intentionally left blank.

List of Tables

This page intentionally left blank.

Glossary

ChalleNGe	National Guard Youth ChalleNGe Program
GED	General Education Development
TABE	Test of Adult Basic Education
WYA	Washington Youth Academy

This page intentionally left blank.

Introduction and Background

This study presents follow-on work involving the analysis of data on changes in youth—both cognitive (e.g., math and language arts) and noncognitive (e.g., ability to follow directions and determination) changes—resulting from their participation in the National Guard Youth ChalleNGe Program (ChalleNGe). CNA's first analysis [1] looked at changes in a class of cadets at the ChalleNGe program in Washington State called the Washington Youth Academy (WYA). This study analyzes data from a second class of cadets from WYA and draws conclusions regarding how participation in ChalleNGe affects youths' cognitive and noncognitive growth. It also looks at the relationship between cognitive and noncognitive measures and the predictive power of noncognitive skills in predicting program completion.

National Guard Youth ChalleNGe Program

The National Guard Youth ChalleNGe Program is designed to provide a second chance to high school dropouts and support for those at risk of dropping out. Eligible youth are ages 16 to 18. The program consists of two components: a 5-month residential portion, followed by a 12-month mentoring phase. ChalleNGe has a quasi-military structure: participants live in barracks, wear military-style uniforms, and perform activities typically associated with military training (e.g., marching, drills, and physical training). However, participants, referred to as *cadets*, participate voluntarily and have no subsequent requirement for military service. The goal of ChalleNGe is to help "young people improve their life skills, education levels, and employment potential" [2].

There are currently 35 ChalleNGe academies operating in 27 states, Puerto Rico, and the District of Columbia. These academies are funded jointly by the Department of Defense and the states. The National Guard Bureau is responsible for management and oversight of ChalleNGe. That said, each site is given discretion in how it structures its program. As a result, the academic goals of the ChalleNGe academies vary. Some seek to have cadets pass the General Education Development (GED) test, while others award alternative high school diplomas. Some ChalleNGe programs provide credit recovery so that cadets can earn high school credits and return to their original high schools after completing the program. There also are some

ChalleNGe academies that are equivalent to high schools and award state-certified high school diplomas.

In addition to providing an academic program, ChalleNGe seeks to instill life skills in the cadets. Toward that end, the core values of ChalleNGe are honor, courage, and commitment. The program also has eight core components: leadership/followership, responsible citizenship, service to community, life-coping skills, physical fitness, health and hygiene, job skills, and academic excellence. All of these core values and components focus cadets toward the changes needed to become productive citizens on completion of the ChalleNGe program.

Some of the goals of ChalleNGe are hard to measure, making an evaluation of the effectiveness of the program difficult. In contrast to academic progress, which can be measured through standardized tests or course completion, some of the core components are heavily dependent on the development of noncognitive skills. The goal of this study is to evaluate changes in cadets' noncognitive skills over the course of the program.

Noncognitive skills

Noncognitive skills are the sets of behaviors, skills, attitudes, and strategies that are not reflected in test scores but play a key role in many areas of life, including career potential, social development, and academic performance. In the literature, noncognitive skills are often referred to as "soft skills." Noncognitive skills can range from study skills, work habits, and time management to individuals' beliefs about their own intelligence, self-control, and persistence. These factors often determine how successfully people manage new environments and meet new academic and social demands [3].

Though noncognitive skills are viewed as important, they often are considered secondary to traditional cognitive skills, such as math and reading proficiency, since the latter can be more easily accessed and measured. Understanding how to improve noncognitive skills is important, however, because—unlike cognitive skills—they are not solely developed in childhood but continue to develop into the young adult years. This means that a program like ChalleNGe has an opportunity to have substantial impact on improving cadets' noncognitive skills. The ChalleNGe program makes concerted efforts to assist students with development of their life skills and other noncognitive measures; this is lacking in the curricula at traditional high schools. For these reasons, we can expect ChalleNGe to have effects on cadets' noncognitive skills that they wouldn't otherwise experience if they remained enrolled in a traditional high school. In a similar program focused on interventions for at-risk minors (albeit younger than those participating in ChalleNGe), the Perry Preschool Project showed long-term success of participants in educational outcomes,

pregnancy rates, criminal behavior, and economic outcomes. These successes are most likely explained by increases in noncognitive skills because the cognitive benefits the participants gained eroded after a short time [4].

Noncognitive skills are important not just because they can be affected well into young adulthood, but also because they have been associated with other positive outcomes. For example, the literature has shown a strong relationship between noncognitive skills and academic success [5]. In addition to the academic benefits, Heckman argues that noncognitive skills are critical in later life, including affecting one's success in the labor market [6]. Other researchers have shown that noncognitive skills are also related to outcomes, such as the probability of arrest/ incarceration and college attendance [4, 7].

The Washington Youth Academy

This study focuses on one particular ChalleNGe academy—the WYA in Bremerton, Washington. The WYA is a particularly well-regarded and efficiently operated program, making it a good candidate for our study. This program operates as a credit recovery system that allows participants to transfer credits back to their home high schools for the coursework they complete at WYA. Each cadet can earn up to 8 credits (approximately 1.3 years of high school credits) to transfer to his or her home high school. The goal is to have cadets return to their home high schools with enough credits to graduate.

Since the noncognitive aspects of ChalleNGe had not previously been studied, in 2013 DOD asked CNA to undertake an evaluation of how cadets' noncognitive skills change over the course of the program. In that study, Wenger and Atkin examined WYA cadets using pre- and post-ChalleNGe surveys and standardized test results. They also evaluated the effectiveness of a new math curriculum based on a facilitated online model in which cadets work independently through modules presented on a computer.[1] The purpose of our current study is to update Wenger and Atkin's analysis using a new dataset from an additional group of ChalleNGe cadets.[2]

[1] The material for the new math curriculum is provided by the Khan Academy—a nonprofit organization with the goal of improving education by "providing a free world-class education for anyone anywhere." The main teaching tool of the Khan Academy is a series of online videos. For more information on the Khan Academy, see [8].

[2] Many of the sections of this report closely follow [1] because this work is the natural extension of the original efforts of Wenger and Atkin. The main way this work is different from the initial study is that the results originate from a new set of ChalleNGe participants.

In the initial study, the authors concluded that, by the end of the ChalleNGe program, noncognitive skills of cadets had improved and gender differences had been eliminated. They also found that the adoption of the online math curriculum led to higher gains in math scores. As a suggestion, the authors recommended that additional data be collected on upcoming classes, to ensure that their findings were not unique to that particular class of cadets. That is the focus of this study.

As a result, during the Fall 2013 class cycle, WYA administered a survey to examine the cadets' noncognitive skills. The survey was identical to the one used in the initial study. It was conducted at both the beginning and end of the program to capture improvements in cadets' noncognitive skills. These data are the basis of our study.

In the next section, we present the data sources that were used in our analysis as well as the methodology used to analyze these data. We then present the results of the study, first focusing on noncognitive changes and then looking at academic improvements. We compare our findings with those reported in [1]. Finally, we end with our conclusions and recommendations.

Data Sources and Methodology[3]

In this analysis, we use several sources of data provided by WYA. First, the program collected cadets' scores on the Test of Adult Basic Education (TABE) at the beginning and the end of the program. In addition, the program collected data indicating which cadets completed ChalleNGe. Finally, all cadets completed a survey that was designed to measure noncognitive skills; they completed the survey twice—once at the conclusion of the initial two weeks (known as pre-ChalleNGe) and again during the last week of classes. In this section, we provide more information on our data sources and how they inform our analysis.

Cognitive skills: TABE scores

Our measure of cognitive skills is created using TABE exam scores, which cadets take at the beginning and the end of ChalleNGe. The TABE was designed for placement of adult learners into appropriate grade-level groups and is often used as an assessment tool in adult education programs that have a focus on GED completion. Each subsection of the TABE is scored to indicate grade level (for example, a score of 9.3 indicates performance at the 3rd month of 9th grade).

Our analysis relies on the four subsections of the TABE, as well as on the total score (formed from averaging subtest scores). The subsections are Math Computation, Applied Math, Reading, and Language. The Math Computation section is made up of computational problems requiring test-takers to perform addition, subtraction, multiplication, and division; to work with percentiles, fractions, and exponents; and to solve basic algebra problems. The Applied Math section comprises word problems, which require the following abilities: chart and table comprehension, basic equation setup, coordinate graphing, an understanding of limited geometry, and application of the concepts of fractions, percentiles, and algebra in the context of word problems. The Language section includes questions on grammar and punctuation, combining

[3] Large portions of this section are taken directly from Wenger and Atkin 2013 [1] because of the similarity of our study structure. Changes are made to the text in areas where our results differ or are presented in other sections.

sentences to preserve their meanings, and some basics of paragraph composition. Finally, the Reading section involves reading passages or detailed charts/tables and answering questions about the content. We chose these four subtests because they represent the core subtests of the TABE. In addition, the ChalleNGe program historically uses these four subtests when reporting test-score data. Finally, of all the TABE subtests, these four are the most similar to the GED.

Noncognitive skills: WYA cadet survey

Our data include several measures of noncognitive skills based on the cadet survey administered by WYA. The cadets completed the survey at the beginning of the program and then completed an identical survey during the last week of the program. The survey included the following measures:

- Grit scale[4]
- Locus-of-control scale[5]
- Efficacy measures to determine cadets' confidence in their math and science abilities[6]
- Time preference—would cadets prefer to be paid $50 today or $100 in 6 months?
- Following directions—cadets were asked to read and follow instructions on a question about why they left their previous high school

The survey's 8-item *grit* scale is designed to measure the respondents' determination or tenacity. For each of these questions, the cadets are presented with a statement and are asked how well it describes them. For example, the survey asks how strongly the cadets agree with the statement, "I finish whatever I begin." The answers range from "Very much like me" to "Not like me at all" in the form of a 5-point Likert

[4] The grit scale was developed by and used with the permission of Dr. Angela Duckworth, Department of Psychology, University of Pennsylvania.

[5] The locus-of-control scale was developed by and used with the permission of Dr. Julian Rotter, Emeritus Professor, Department of Psychology, University of Connecticut. In both Rotter's and our work, an internal locus of control is considered to be a positive attribute.

[6] Efficacy scales were adapted from Middle and High School STEM-Student Survey, 2012, Raleigh, North Carolina, and used by permission of the Friday Institute for Educational Innovation, North Carolina State University.

(rating) scale. The grit score is calculated by awarding points for stated determination; for example, one statement is, "I am a hard worker," and another is "I often set one goal but later choose to pursue a different goal." For the first statement, cadets received 5 points for selecting "Very much like me" and decreasing numbers of points down to 1 point for "Not at all like me." For the second statement, cadets received 1 point for choosing "Very much like me" and increasing numbers of points up to 5 points for "Not at all like me." Total scores range from 8 to 40 with higher scores indicating higher levels of determination, or grit.

Locus-of-control measures the extent to which a person believes that his or her own actions (versus random factors or other powers) determine outcomes. Essentially, the scale measures the extent to which respondents believe that they can control their lives. Those who believe that their own actions have consequences are designated as "internal"; those who believe that other factors determine outcomes are termed "external." For each question, the respondent chooses which of two statements best describes his or her beliefs/feelings. Respondents receive 1 point each time they choose a statement indicating that they have control over situations; the score ranges from 0 (completely external, failing to see a relationship between their own actions and consequences/reactions) to 13 (completely internal, giving no explanatory power to luck). We consider an internal locus-of-control to be preferable (and therefore assign it a higher value); people with an internal locus-of-control are more likely to take actions that will result in positive consequences or rewards because they see a direct correlation between outcomes and their own behaviors. Conversely, those with an external locus-of-control will be less likely to take responsibility for any negative outcomes that occur in their lives; they will therefore not be likely to adjust their behaviors accordingly.

Efficacy is measured using a 5-point Likert scale of responses to a series of statements about the cadet's attitude toward, and confidence in, math and science. We calculate math and science efficacy separately. In each case, the efficacy score is determined by awarding points for responses that exhibit a positive attitude or confidence in the subject. Thus, cadets who select "Strongly agree" for such statements as "I know I can do well in science" receive 5 points, as do cadets who select "Strongly disagree" for such statements as "I can handle most subjects well, but I cannot do a good job in science." Each efficacy score indicates the average response on the Likert scale with higher scores indicating higher efficacy. Scores range from 1 to 5.

Time preference is the fourth measure of noncognitive skills and is captured by a simple question asking whether the cadet would prefer to be paid $50 today or twice that in 6 months. Indicating a preference for $100 in 6 months suggests a level of determination, planning, and self-control.

Following directions is the final measure. Cadets are asked why they left their previous high school. They are presented with a variety of reasons and are instructed

to mark all that apply as well as to circle the most important reason. All cadets marked at least one reason. We considered those who also circled a reason to have followed the directions and those who did not circle a reason to have not followed the directions.[7]

A total of 143 cadets filled out the initial survey. During the classroom phase, 19 cadets left the program; thus, 124 cadets completed the program.[8] Due to a medical absence, we do not have a final survey on one cadet. Therefore, we have 123 complete, matched surveys (including pre- and post-ChalleNGe information). In a few cases, cadets skipped questions or sections of the survey, but, overall, cadets answered the vast majority of the questions on the pre- and post-ChalleNGe surveys. In evaluating the measures, we present the most complete information possible and use all partial information provided in the survey to the fullest extent possible.

[7] For a comprehensive review of each of these noncognitive measures and a more in-depth discussion of why they are viewed as beneficial to development, see [9].

[8] In terms of percentages, we experienced the same attrition as in the first iteration of this study. Of the 152 cadets who entered the program in January 2013, 133 completed the program, for an attrition rate of 13 percent.

Results

In this section, we present results from our analysis of the Fall 2013 WYA cycle. Initially, we focus on the survey results of the cadets' noncognitive skills, providing descriptive statistics and explaining how they change during the program. We also examine the progress they made in cognitive skills and then analyze the relationship between noncognitive skills and program outcomes. Using TABE data, we determine if initial TABE scores, in addition to initial noncognitive scores, can be used to predict ChalleNGe completion and how noncognitive skills influence TABE scores.

Noncognitive skills

As previously discussed, we ascertain cadets' level of noncognitive skills from a survey. There are three comparison groups whose survey results we analyze:

1. The pre-ChalleNGe survey of all cadets[9]
2. The pre-ChalleNGe survey of cadets who complete ChalleNGe
3. The post-ChalleNGe survey of cadets who complete ChalleNGe

These groups are meaningful because they allow us to establish the two sets of comparisons of primary interest in this study. The first is to compare the initial noncognitive skills of cadets who start ChalleNGe, but do not finish, with those of cadets who complete ChalleNGe. This comparison allows us to analyze whether there are statistically significant differences in the noncognitive skills of those who complete ChalleNGe versus those who do not. The second is to compare the initial noncognitive skills of cadets entering ChalleNGe with the final noncognitive skills of these same cadets, once they graduate. This comparison provides an understanding of whether cadets who complete ChalleNGe experience an improvement in their noncognitive skills as a result of their participation in the program.

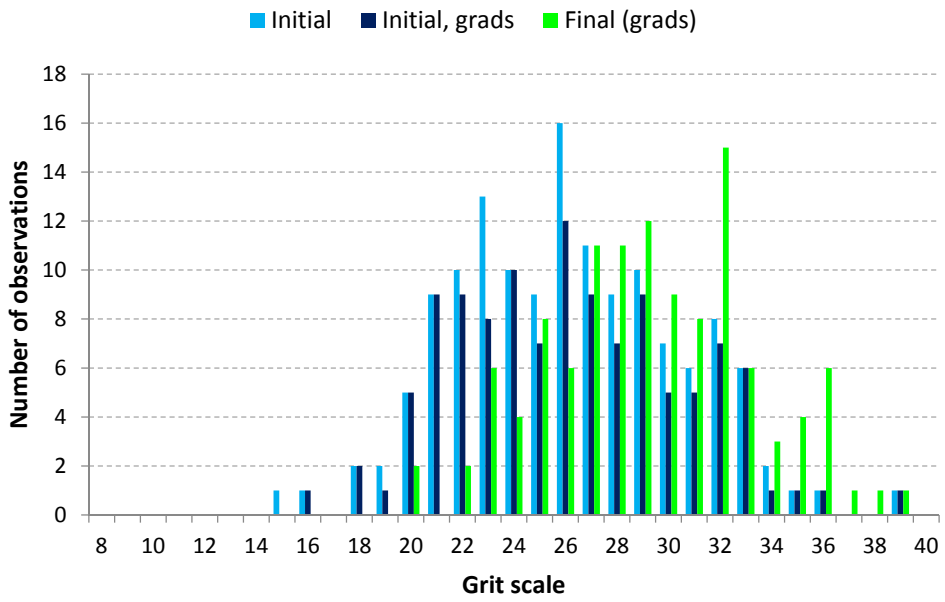
[9] An alternative to defining this group as all cadets who took the pre-ChalleNGe survey is to define it as those cadets who do not finish ChalleNGe. Although analytically interesting, this alternative group would contain only 19 cadets and would therefore be too small to use as the basis for making statistical conclusions.

Descriptive statistics

Before exploring the comparison of the three groups of survey results, we provide descriptive statistics for each of the metrics we analyze in the cadet survey. The following figures present the score distributions for the grit, locus-of-control, and efficacy measures. Each figure presents both pre- and post-ChalleNGe scores for all graduates, along with pre-ChalleNGe scores for all cadets. Light blue and dark blue bars, respectively, show the initial score distributions for all cadets and for cadets who ultimately graduate. Green bars represent the distribution of final grit scores (no final scores are available for cadets who do not complete ChalleNGe, which is why the final grit scores are shown for graduates only).

Figure 1 shows the distribution of cadets' grit scores. The modal initial grit score is 26 both for cadets who start ChalleNGe and complete the program and for those that do not complete the program.[10]

Figure 1. Cadets' grit-score distributions, pre- and post-ChalleNGe

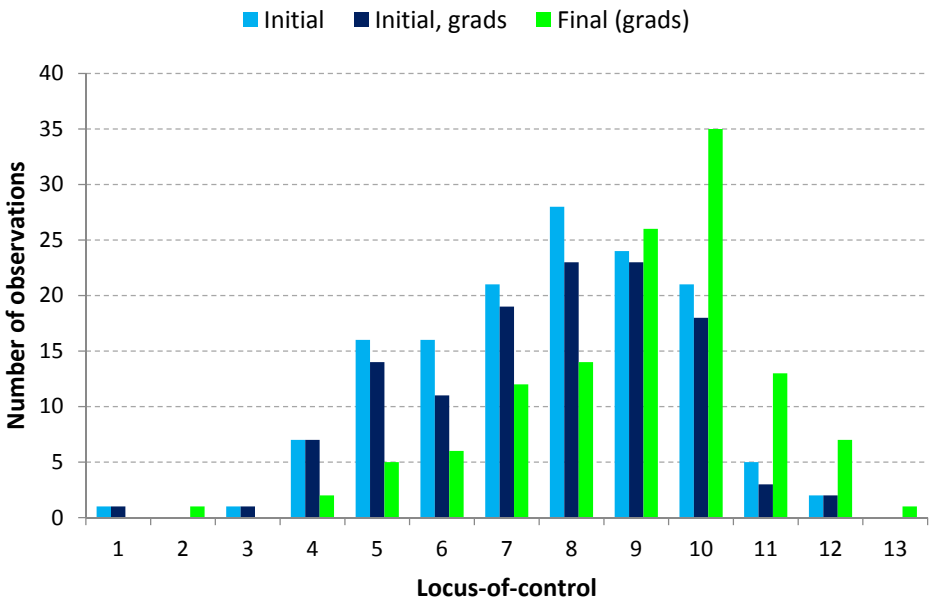


[10] We use the mode to indicate average behavior for each metric and chose it over, for example, the mean, because the mode is the most visually recognizable measure of central tendency in these figures. In all cases in this report, the mean and the median are similar to the mode because of the unimodal distributions, as illustrated in the figures.

By the end of ChalleNGe, however, cadets have more grit. This is seen in the shift toward the right of the green bars in Figure 1. The mode of the final grit score of the graduating cadets is 32. This improvement suggests that cadets are becoming more determined (i.e., have higher grit) as a result of the ChalleNGe program.

Figure 2 illustrates the distribution of cadets' locus-of-control. The mode of the initial locus-of-control distribution of all cadets is a score of 8. Similarly, the mode of the initial locus-of-control distribution for cadets who ultimately graduate is 8 and 9.[11] However, there is a rightward shift in the distribution of the final locus-of-control. Specifically, the modal locus-of-control score for cadets who graduate from ChalleNGe is 10. This means that these cadets become more internal (i.e., they see a closer link between their actions and consequences, thus giving less credence to luck determining outcomes) as a result of their participation in the ChalleNGe program.

Figure 2. Cadets' locus-of-control distributions, pre- and post-ChalleNGe



[11] Typically, the mode is represented by a single value, but in this case there is a tie for the value with the most observations. Therefore, there are two modes.

Figure 3 provides the distribution of cadets' math efficacy scores. The mode of the initial math efficacy distribution for all cadets and graduates is a score between 2 and 3. Once again, we notice a rightward shift of the final math efficacy scores, represented by the green bars in Figure 3. In this case, there is no change in the mode, but distribution still shifts to the right because of an increase in the number of cadets who score in the highest two ranges of the figure. This indicates that cadets who complete ChalleNGe experience an increase in their math-skill confidence.

Figure 3. Distribution of math efficacy for cadets pre- and post-ChalleNGe

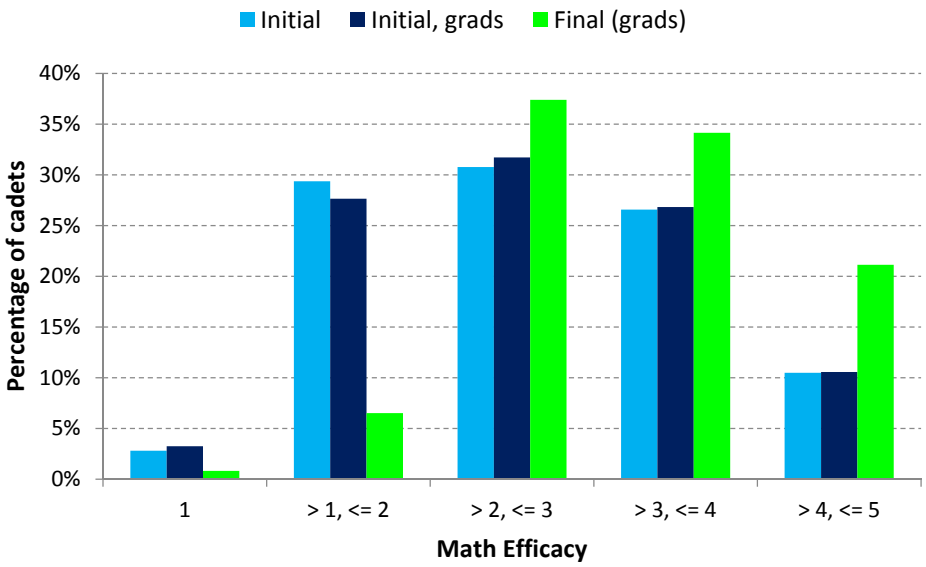
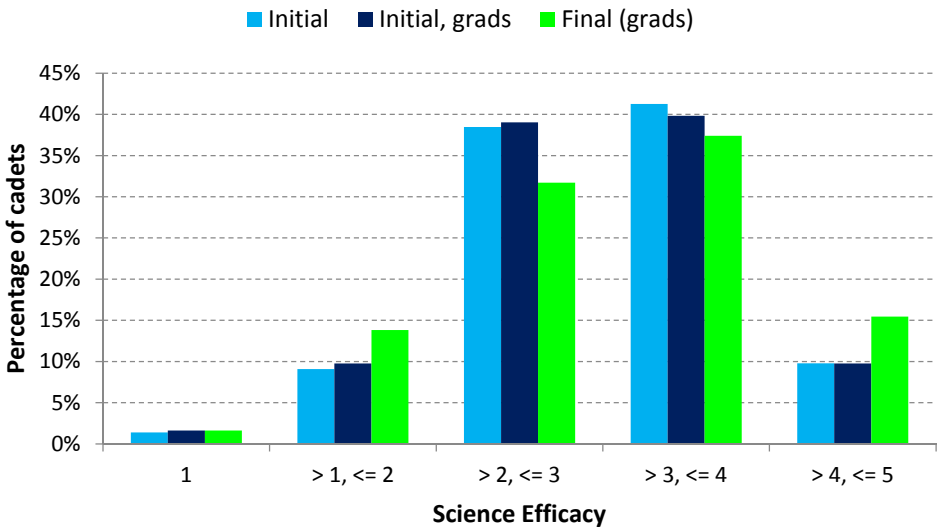


Figure 4 shows the distribution of cadets' science efficacy scores. The mode of all three series shown (initial cadets' scores, initial cadets' scores for those who ultimately graduate, and final scores of graduates) is between 3 and 4. In contrast to the previous measures presented, there is no clear shift or improvement in pre- and post-ChalleNGe scores for science efficacy. This may be a consequence of cadets' relatively high initial science efficacy (the majority of scores are above 3), and the scale we are using has a maximum value of 5. Thus, the maximum improvement that the average cadet could experience in science efficacy is less than 2.

The final two noncognitive measures are time preference and following directions. In the initial survey, 52 percent of cadets opted to receive $100 in 6 months (as opposed to $50 today). This percentage improved to 75 percent in the final survey. This 44-percent increase indicates that cadets, after completing ChalleNGe, become more willing to accept delayed gratification. When responding to the portion of the survey designed to evaluate how well cadets follow directions, 14 percent of cadets

did so in the initial survey. In the final survey, 28 percent followed directions. This 100-percent increase clearly indicates that cadets' ability to read and follow directions improves as a consequence of their ChalleNGe participation.

Figure 4. Distribution of science efficacy for cadets pre- and post-ChalleNGe



Comparison of pre- and post-ChalleNGe scores

The survey results are presented in Table 1. Each value represents the average score for the group of cadets. We include the initial scores for all cadets who started ChalleNGe as well as the initial scores for only those who graduated from the program. We also provide the final scores for those cadets who graduated.

Table 1 illustrates two main points, which are consistent with the analysis in [1]. The first is that, among the cadets who ultimately graduated from ChalleNGe, noncognitive skills improved on average. This can be seen by comparing the last two columns of Table 1. Cadets who finished the program had statistically significantly higher scores in all noncognitive measures, except science efficacy, where no statistically significant progress is made. The results show that cadets who complete ChalleNGe have more grit or fortitude, are more internal (meaning they are more likely to believe that *they* have greater control over their destiny), and have greater efficacy or confidence in math. A higher percentage of cadets also show willingness to choose to receive $100 in 6 months versus $50 today, which indicates an understanding of delayed gratification and greater self-control. And, finally, a higher percentage of cadets are more likely to follow directions than they were at the beginning of the program. For a comprehensive discussion of the importance of

noncognitive skills, not only in the educational environment but also in determining posteducation outcomes, see [10].

Table 1. Noncognitive measures, before and after ChalleNGe[a]

Noncognitive measure	Initial score		Final scores, graduates
	All cadets	Graduates	
Grit score	26.1	26.2	29.3**
Math efficacy	2.7	2.7	3.2**
Science efficacy	3.0	3.0	3.1
Locus-of-control (internal)	7.6	7.6	8.9**
Chose $100 in 6 months (%)	51.8	51.7	74.2**
Followed directions (%)	14.0	15.1	27.7*

a. Sample sizes vary for the various metrics based on the number of respondents in the survey. In all cases, the variation is minimal and does not affect interpretation of the results.

** Differences between graduates' initial and final scores are statistically significant at the 1-percent level (likelihood of occurring by chance less than 1 in 100).

* Differences between graduates' initial and final score are statistically significant at the 5-percent level (likelihood of occurring by chance less than 1 in 20).

The second point that Table 1 illustrates is that the initial measures of noncognitive skills are not good predictors of which cadets will complete the ChalleNGe program. Specifically, when comparing the first and second columns of the table (where initial scores of the cadets who ultimately graduate are compared with those of all cadets who start ChalleNGe), we see that the average scores are almost identical in all noncognitive measures, except for following directions.

A slightly higher percentage of graduates than the percentage of all cadets followed directions on the pre-ChalleNGe survey. Table 1 shows that 15 percent of cadets who graduate follow directions in the initial survey, compared with 14 percent of all cadets. This small difference suggests that this measure is not predictive of success. If it were, we would expect to observe sufficiently higher scores for graduates as compared with all initial cadets.

These results mirror the findings in [1], showing that ChalleNGe is having a substantial, positive impact on cadets' noncognitive abilities and that cadets' initial noncognitive skills cannot be used to predict ChalleNGe success.

Gender comparison

Next, we examine these same noncognitive skills by gender. These results will enable program directors to determine whether and how their approach could be tailored for female versus male cadets. Table 2 shows the performance of each group on noncognitive measures.

Table 2. Initial and final scores on noncognitive measures, by gender[a]

Noncognitive measure	Initial score of cadets		Final score of cadets	
	Female	Male	Female	Male
Grit score	25.2	26.6	27.1	30.3**
Math efficacy	2.3	2.9**	2.8	3.1**
Science efficacy	2.6	3.2**	2.7	3.3**
Locus-of-control (internal)	7.1	7.8*	8.5	9.2
Chose $100 in 6 months (%)	58.3	48.2	66.7	77.9
Followed directions (%)	18.9	12.8	27.0	28.0

[a.] Sample sizes vary for the various metrics based on the number of respondents in the survey. In all cases, the variation is minimal and does not affect interpretation of the results.

** Differences between men and women are statistically significant at the 1-percent level (likelihood of occurring by chance less than 1 in 100).

* Differences between men and women are statistically significant at the 5-percent level (likelihood of occurring by chance less than 1 in 20).

The results in Table 2 show that male cadets begin ChalleNGe with significantly higher scores in math and science efficacy. This is consistent with CNA's other research finding that female students hold lower science and mathematical self-efficacy than their male counterparts [11]. Male cadets also are initially more internal than female cadets. This means that male cadets believe they have greater control over outcomes in their lives than female cadets do.

In the final scores, there are significant differences between male and female cadets in the areas of math and science efficacy, as there were in the initial scores. Also, a significant difference between genders exists for the final grit score. The initial gender difference in locus-of-control is not evident in the final scores. These gender differences may indicate that male and female cadets might require different approaches to the noncognitive aspects of the ChalleNGe curriculum.

Recall that the primary goals of the ChalleNGe program are twofold, including both noncognitive and cognitive skills. Thus, although our primary focus is on the development of noncognitive skills, we also evaluated the ChalleNGe program's

impact on cadets' cognitive skills. We do this by analyzing changes in TABE test scores from the beginning of the program to the end.

Cognitive skills

The cognitive skills of the cadets were measured with TABE exam scores. Table 3 presents average scores for the four TABE subtests, both on arrival at ChalleNGe and at graduation (for those who completed the program).

Table 3. Cognitive measures[a]

TABE subtest	Initial score		Final scores, graduates
	All cadets	Graduates	
Reading	6.8	6.8	10.3**
Language	6.8	6.9	9.4**
Math Computation	7.0	6.9	9.1**
Applied Math	9.3	9.2	11.1**

a. Sample sizes vary for the various metrics based on the number of respondents in the survey. In all cases, the variation is minimal and does not affect interpretation of the results.

** Differences between initial and final score among graduates are statistically significant at the 1-percent level (likelihood of occurring by chance less than 1 in 100).

In Table 3, we observe that the average TABE score for all subtests is about 7, with the exception of applied math. This means that the average cadet is entering ChalleNGe at a 7th grade level in reading, language, and math computation. The outlier is applied math where cadets start with an average score of 9.2, or at the 9th grade level. (Recall that TABE scores are grade-level equivalents, so a value of 9.2 means that a cadet is performing at the level of the 2nd month of 9th grade.) This suggests that cadets are much more proficient in applied math than the other TABE components.

Table 3 also illustrates that ChalleNGe graduates make significant progress in *all* cognitive measures. The final scores, shown in the last column of the table, represent nearly a two-grade-level improvement in all TABE subtests. This average two-grade-level improvement across subtests was found in the previous study as well. The largest improvement occurs with the Reading subtest, where graduating cadets improve 3.5 grade levels, on average, from their initial scores (in the previous study, the average reading improvement was 1.7 years). This is a remarkable achievement over the course of a 5.5-month program. This is especially true when considering that the average cadet gains only 2 years in TABE scores during the course of the

program, when accounting for all WYA TABE data from 2009 through 2013. The average reading gain, for example, is only 1.7 years among all cadets at WYA during this time period [1]. To contrast this gain, the smallest subtest improvement for WYA cadets occurs in the area of applied math, where cadets improve, on average, by 1.9 grade levels (in the previous study, applied math and reading both experienced the lowest gains, at 1.7 grade levels). Nonetheless, this is a statistically significant improvement which occurred despite the fact that (1) the initial applied math score is the highest among all subtests and (2) the scores cannot be higher than 12.9.

Table 3 also shows that there are no significant differences between the initial TABE scores of all cadets and the cadets who eventually graduate. This is represented in the first two columns of Table 3, where all values are within 0.1 grade level. The closeness of the initial TABE scores of all cadets who started ChalleNGe and those who eventually graduate leads us to conclude that initial TABE scores are not a good predictor of program completion.

Next we examine these same TABE test results by gender. Table 4 shows the performance of female and male cadets in each of the TABE subtests.

Table 4. Initial and final scores on cognitive measures, by gender[a]

TABE subtest	Initial score of cadets		Final score of cadets	
	Female	Male	Female	Male
Reading	6.3	7.0	9.8	10.5
Language	6.9	6.9	8.6	9.7*
Math Computation	5.6	7.5**	7.8	9.7**
Applied Math	8.3	9.6*	10.2	11.6**

[a.] Sample sizes vary for the various metrics based on the number of respondents in the survey. In all cases, the variation is minimal and does not affect interpretation of the results.

** Differences between men and women are statistically significant at the 1-percent level (likelihood of occurring by chance less than 1 in 100).

* Differences between men and women are statistically significant at the 5-percent level (likelihood of occurring by chance less than 1 in 20).

Consistent with our noncognitive findings, the results in Table 4 show that male cadets begin ChalleNGe with significantly higher scores in both Math subtests. Female cadets enter with average scores of 5.6 and 8.3 in math computation and applied math, respectively. There is little difference by gender in average reading and language scores; both male and female cadets average a 6.5 level for reading and a 6.9 level for language.

In the final TABE scores, shown in the last two columns of Table 4, there remains a significant gender difference in math computation: male cadets have an average final

score of 9.7, compared with 7.8 for female cadets. This male-female gap is even greater in language and applied math. Whereas there were no significant differences in the average scores of male and female cadets in the initial language test, a significant difference develops in the final test.

In summary, we find improvements in both noncognitive and cognitive skills, in addition to statistically significant differences by gender. With respect to noncognitive skills, the largest areas of improvement are in choosing to receive $100 in 6 months (an indication of understanding the value of delayed gratification) and in following directions. The results of the TABE show that there is statistically significant evidence that TABE scores improve for all cadets on average. This suggests that ChalleNGe is effectively improving both the noncognitive and cognitive abilities of its cadets.

Relationship between cognitive and noncognitive skills

Having separately analyzed cadets' improvements in cognitive and noncognitive skills throughout the ChalleNGe program, we now determine how these two skill types are related, if at all. As an analytical exercise, we attempt to predict cadets' final math TABE scores as a function of their initial math TABE scores *and* their initial math efficacy.[12] This will inform us as to whether, in the case of final math scores, the initial cognitive or the initial noncognitive skill is more important.

We found that creating strata of the final math scores led to the best-fitting model. We therefore estimate two separate models: one for cadets with relatively low test scores (below 6) and one for those with relatively high test scores (greater than or equal to 6). Each equation included the initial TABE math score and the initial math efficacy. The results indicate that, in the low-test-score group, both the initial math score and the initial math efficacy have positive and statistically significant impacts on the final math score, for both genders. This suggests that both the initial math score and initial math efficacy are important predictors of the final TABE math score for low-achieving math students. According to the recommendation in [1], these results suggest that it is especially important to work to educate all cadets on the

[12] We conduct this analysis on final math scores only (in lieu of also predicting final reading scores) because there was a recent change in the math curriculum prior to data collection. Specifically, WYA began using Khan Academy for math instruction in 2009. Thus, the ChalleNGe staff was particularly interested in the impact of initial math TABE scores combined with initial math efficacy on final math scores.

importance of improving their mathematical abilities, and to increase their confidence about their math skills, before beginning the classroom portions of ChalleNGe.

For the higher initial math score group, the results are somewhat different. In this case, only the initial math TABE score is a statistically significant predictor, and this finding holds for male cadets only. Thus, as final math scores increase, it becomes more difficult to predict these scores. These findings suggest that initial math efficacy is much more important in predicting final math scores for the low-test-score group than the high-test-score group. Of course, the test-score group for a particular cadet is not known until the end of the program, when final TABE results are collected. Thus, efforts to improve cadets' perceived math efficacy should be directed at all cadets. There will be no harm for those cadets who will ultimately be higher scoring on the final TABE, and significant gains can be realized for those who will fall into the lower scoring group. Thus, there are no "costs" to directing such efforts at all cadets.

Predictive power of noncognitive measures

While it is informative to understand on which cognitive and noncognitive measures cadets experience the greatest improvements during the ChalleNGe program, an investigation into the relationship between noncognitive measures and ChalleNGe program completion is also important. If we were able to identify particular noncognitive measures that are positively (or negatively) associated with program completion, WYA could select cadets on these characteristics (when concerned with their attrition rates) and/or spend additional program time working on improving these noncognitive skills, to make program completion more likely.

Thus, we now present a simple model of ChalleNGe completion as a function of both cognitive and noncognitive measures. Specifically, we explain the results of a regression model in which the dependent variable is dichotomous: cadets either complete ChalleNGe or they do not. We use a logistic (logit) regression model.

Since we are working with a relatively small dataset, we are only able to estimate only a number of simple equations, which include just a few variables. We have already shown that there is little difference between the initial scores of all cadets and those who graduate (and thus most noncognitive and cognitive measures will not have a significant impact on a cadet's ability to complete ChalleNGe), so our estimation is not hindered by the exclusion of some initial scores.

Our final model includes initial grit scores and the initial cognitive measures from the TABE. Other noncognitive measures were not included because they have little impact on program completion. Because of missing information on some pre- and

post-ChalleNGe surveys, there was not a perfect match (i.e., a cadet may have answered the grit question on the pre-ChalleNGe survey but not on the post-ChalleNGe survey). Taking all of this into account, caution should be used when interpreting these results on program completion. Specifically, although we report average initial and final scores for each noncognitive measure, we are not able to calculate the improvement for every cadet, since some cadets did not answer all questions on either the initial or the final survey.

Our estimations revealed that no single variable had enough explanatory power to be statistically significant in predicting ChalleNGe program graduation. However, the language scores were the closest to being statistically significant and have the proper sign for the coefficient. Similarly, the reading scores had the proper sign for its coefficient, but they were not statistically significant. We suspect that the statistical insignificance of these measures is a result of our sample size, as opposed to indicating that there is no meaningful relationship between these scores and the probability of completion. The relationship between reading scores and noncompletion was established in [1] (and was statistically significant); therefore, it is not surprising that this relationship appears again. The regressions also reveal that initial reading scores and language scores have positive effects on the likelihood of completion, although these results are also statistically insignificant.[13] Finally, the relationship between initial grit and noncompletion, although insignificant, also had the proper sign.

[13] Complete regression results are provided in the appendix.

Conclusions and Recommendations

In this report, we have extended the work of previous CNA research on the success of the WYA ChalleNGe program in achieving both noncognitive and cognitive gains for the cadets. In doing so, we analyzed data from the most recent data collection effort: the Fall 2013 class of cadets. Our results mirror the results of the initial study in many ways, while exploring the relationships that are brought forth by this new class of cadets.

Our findings suggest that the WYA ChalleNGe program continues to have a substantial impact on cadets' noncognitive skills. At the conclusion of the ChalleNGe program, our data show that cadets have more grit (determination), a greater internal locus-of-control, a greater ability to follow directions, a greater willingness to take $100 in 6 months in lieu of $50 today (delayed gratification), and higher math efficacy. The only noncognitive measure that does *not* improve significantly is science efficacy. This appears to result from the fact that over 50 percent of cadets start with a science efficacy equal to or above 3 on a fixed scale of 1 to 5. Thus, there is not much room for improvement in science efficacy on average. Based on our research on the WYA ChalleNGe program to date, the program appears to improve cadets' noncognitive skills.

Along with looking for signs of improvement in noncognitive skills, we also analyzed improvements in cadets' cognitive skills over the course of the program. We find that statistically significant improvements are made in all four TABE subtests that we analyzed. From induction to graduation, cadets improve, on average, by at least 2 grade levels in every subtest (Reading, Language, Math Computation, and Applied Math). For the Reading subtest, this particular cadet class experienced average improvements of 3.5 grade levels, which is twice the average improvement from 2009 to 2013. In addition, we also find no significant differences between the initial TABE scores of all cadets and the initial TABE scores of those cadets who eventually graduate. This indicates that initial cognitive abilities are *not* predictive of program completion, and it suggests that ChalleNGe has made somewhat of a difference for these cadets. Had ChalleNGe not been effective in influencing these cadets in some way, we would expect those with lower scores to be less likely to complete the program.

We suspect that the positive changes that occur in cadets' cognitive and noncognitive skills will be long lasting; at this time, however, we have no quantitative metrics to

determine if this is true. We therefore suggest that WYA, and other ChalleNGe academies, strongly consider a longitudinal study of cadet performance. This would allow cadets to be tracked as they reintegrate into their home high schools (where applicable), and eventually into their postsecondary schooling and career fields. These data would provide the necessary information for analysis regarding whether the noncognitive skills gained in ChalleNGe last beyond the program's end, as recent literature suggests [4].

We are unable to find any noncognitive measures that strongly predict program completion. The initial grit score shows promise but is not statistically significant in our model. Similarly, the cognitive factors of initial reading and language scores also have the correct signs in our model but are not statistically significant. As previously discussed, this may be more a reflection of our small sample size than of the inherent relationship (or lack thereof) between these scores and program completion.

Finally, we are able to establish a relationship between math efficacy, initial math score, and the final math score. This required us to estimate separate equations for cadets with low initial math scores and those with high initial math scores. We also estimated these models separately by gender. This allowed us to determine that both the initial math score and initial math efficacy are strong predictors of the final math score, for both male and female cadets, for those with initial math scores below 6.0. However, for those cadets with initial math scores greater than or equal to 6.0, only the initial math score is a significant predictor of the final math score, and this finding only holds for male cadets. These findings suggest that initial math efficacy is much more important in predicting final math scores for the low-test-score group than the high-test-score group. Since cadets' final math scores are not known until the end of ChalleNGe, we recommend that efforts to improve cadets' perceived math efficacy be increased and directed at *all* cadets.

In addition to this final TABE score finding, a number of our other findings also differ by gender. First, male and female cadets begin ChalleNGe with significant differences in their cognitive and noncognitive skills; The young men have higher initial math and science efficacy, have greater initial locus-of-control, and have higher initial TABE math scores (on both the Math Computation and Applied Math subtests). Second, their scores also are statistically significantly different at the end of ChalleNGe. In terms of noncognitive skills, male graduates have higher math efficacy, science efficacy, and grit. In terms of final cognitive skills, males have statistically higher final TABE scores in the Math Computation, Applied Math, and Language subtests. As we previously suggested, these gender differences could suggest that gender-tailored approaches would be appropriate within the ChalleNGe curriculum.

As indicated, our results on the noncognitive measures mirror those in Wenger and Atkin [1] in several areas. First, among the cadets who ultimately graduated from ChalleNGe, noncognitive skills improved on average in both studies. Cadets who finished the program had statistically significantly higher scores in all noncognitive

measures, except science efficacy where no statistically significant progress was found in our most recent work. Second, both efforts found that initial measures of noncognitive skills are not good predictors of which cadets will complete the ChalleNGe program. Both studies also found gender differences in the noncognitive measures. Specifically, female cadets began the program with lower measures of efficacy (in science and math) and were less internal than male cadets. Finally, with respect to cognitive changes, both studies concluded that the TABE reading score had explanatory power over program completion. Cadets with lower reading scores were less likely to complete the program.

Appendix A: Regression Results

The following tables provide the detailed regression estimates discussed in the main body of the paper. Table 5 shows that initial reading and initial grit are the two most important factors in predicting ChalleNGe noncompletion. They are both positively associated with program completion (and negatively associated with noncompletion). Also, initial language is negatively associated with noncompletion, but to a lesser degree. Ultimately, the model is not a very good fit.

Table 5. Predictors of ChalleNGe noncompletion[a, b]

Variable	Coefficient	Standard error
Initial grit	-0.226	0.06
Initial math	0.761	0.147
Initial reading	-0.242	0.111
Initial language	-0.155	0.118
Initial applied math	0.0644	0.164
Constant	-1.54	1.93

[a.] Regression includes 139 observations (all cadets with complete matched test score and survey data). Pseudo R-squared = 0.02. Initial grit is measured by grit scale, developed by Dr. Angela Duckworth. Initial math, reading, language, and applied math are measured, respectively, by the TABE Math Computation subtest, the TABE Reading subtest, the TABE Language subtest, and the TABE Applied Math subtest.

[b.] None of these initial scores are statistically significant predictors of noncompletion.

Table 6 and Table 7 include results of a simple linear regression of final math scores as a function of initial math score and initial math efficacy. We stratify the population into two subpopulations, one for relatively low initial math scores (< 6.0) and the other for relatively high initial math scores (>= 6.0). Each equation is also run separately for male and female cadets. When the initial math score is less than 6.0, both initial math efficacy and initial math score are positively associated with the final math score for both male and female cadets. However, when the initial math score is greater than or equal to 6.0, only the initial math score is positively associated with the final math score, and this is true for male cadets only.

Table 6. Predictors of final math TABE scores for those with low (< 6.0) initial math TABE scores[a]

	Female cadets		Male cadets	
Variable	Coefficient	Standard error	Coefficient	Standard error
Initial math score	1.024**	0.191	0.689**	0.231
Initial math efficacy	0.073*	0.266	0.771*	0.15
Constant	2.066*	0.765	2.29	1.40

[a.] Regression includes 33 observations on male cadets and 20 on female cadets (all cadets with complete matched test score and survey data). Adjusted R-squared = 0.26 for men and 0.66 for women. Math efficacy is measured by a scale developed by Friday Institute, North Carolina State University. Initial math is measured by the TABE Math subtest.

** Indicates that coefficient is significant at the 1-percent level or better and, thus, is likely to occur by chance fewer than 1 time in 100.

* Indicates that coefficient is significant at the 5-percent level or better and, thus, is likely to occur by chance fewer than 1 time in 20.

Table 7. Predictors of final math TABE scores for those with high (>=6.0) initial math TABE scores[a]

	Female Cadets		Male Cadets	
Variable	Coefficient	Standard error	Coefficient	Standard error
Initial math score	0.610	0.590	0.421**	0.148
Initial math efficacy	0.381	0.556	0.486	0.315
Constant	3.954	3.808	5.714**	1.23

[a.] Regression includes 53 observations on male cadets and 17 on female cadets (all cadets with complete matched test score and survey data). Adjusted R-squared = 0.27 for men and 0.18 for women. Math efficacy is measured by a scale developed by Friday Institute, North Carolina State University. Initial math is measured by the TABE Math subtest.

** Indicates that coefficient is significant at the 1-percent level or better and, thus, is likely to occur by chance fewer than 1 time in 100.

References

[1] Wenger, Jennie W., and Jennifer R. Atkin. 2013. *Cognitive and Noncognitive Changes From Participation in National Guard Youth ChalleNGe.* CNA Corporation. DRM-2013-U-005201.

[2] National Guard Youth Foundation. 2014. "About Youth ChalleNGe." National Guard Youth Foundation. Sep. 17, 2014. http://www.ngyf.org/about_youth_challenge.html.

[3] Farkas, G. 2003. "Cognitive Skills and Noncognitive Traits and Behaviors in Stratification Processes." *Annual Review of Sociology* 29: 541-562.

[4] Heckman, James J., Seong Hyeok Moon, Rodrigo Pinto, Peter A. Savelyev, and Adam Yavitz. 2010. "The Rate of Return to the HighScope Perry Preschool Program." *Journal of Public Economics* 92 (1-2): 114-128.

[5] Farrington, C. A., M. Roderick, E. Allensworth, J. Nagaoka, T. S. Keyes, D. W. Johnson, and N. O. Beechum. 2012. *Teaching Adolescents To Become Learners. The Role of Noncognitive Factors in Shaping School Performance: A Critical Literature Review.* University of Chicago Consortium on Chicago School Research.

[6] Heckman, J. J. 2008. "Schools, Skills, and Synapses." *Economic Inquiry* 46 (3): 289-324.

[7] Bowles, Samuel, Herbert Gintis, and Melissa Osbourne. 2001. "The Determinants of Earnings: A Behavioral Approach." *Journal of Economic Literature* 39 (4): 1137-1167.

[8] Khan Academy. 2014. Khan Academy Homepage. Sep. 18, 2014. https://www.khanacademy.org/.

[9] Wenger, Jennie, Cathleen McHugh, Seema Sayala, and Robert Shuford. 2008. *Variation in Participants and Policies Across ChalleNGe Programs.* CNA Corporation. CRM D0016643.A2/Final.

[10] Rosen, Jeffrey A., Elizabeth J. Glennie, Ben W. Dalton, Jean M. Lennon, and Robert N. Bozick. 2010. *Noncognitive Skills in the Classroom: New Perspectives on Educational Research.* RTI International. BK-0004-1009. Accessed Oct. 6, 2014. http://www.rti.org/pubs/bk-0004-1009-rosen.pdf.

[11] Louis, Rachel A., and Jean M. Mistele. 2012. "The Difference in Scores and Self-Efficacy by Student Gender in Mathematics and Science." *International Journal of Science and Mathematics Education* 10 (5): 1163-1190. doi: 10.1007/s10763-011-9325-9.

The CNA Corporation

This report was written by CNA Corporation's Resource Analysis Division (RAD).

RAD provides analytical services—through empirical research, modeling, and simulation—to help develop, evaluate, and implement policies, practices, and programs that make people, budgets, and assets more effective and efficient. Major areas of research include energy and environment; manpower management; acquisition and cost; infrastructure; and military readiness.

DRM-2014-U-008846-Final-2

CNA Corporation is a not-for-profit research organization
that serves the public interest by providing
in-depth analysis and result-oriented solutions
to help government leaders choose
the best course of action
in setting policy and managing operations.

Nobody gets closer—
to the people, to the data, to the problem.

CNA
ANALYSIS & SOLUTIONS

www.cna.org • 703-824-2000

3003 Washington Boulevard, Arlington, VA 22201